**Welcome to DialogClassic Web(tm)**

```
 Dialog level 01.04.26D
Reconnected in file OS 26may01 10:47:29
HILIGHT set on as ' '
KWIC is set to 50.
705CORE1 is set ON as an alias for 16,160,148,621,275.
705CORE2 is set ON as an alias for 15,9,623,810,624,636,813,20.
705NFT is set ON as an alias for 77,35,583,65,2,233,99.
705NEWS is set ON as an alias for 473,474,475.
705SOFT is set ON as an alias for 278,634,256.
705PATF is set ON as an alias for 348,349.
705ADS is set ON as an alias for 635,570,PAPERSMJ,PAPERSEU.
 * * * File 278 is currently unavailable. * * *
>>>         278 is temporarily unavailable
>>>1 of the specified files is not available

SYSTEM:OS  - DIALOG OneSearch
  File  16:Gale Group PROMT(R)  1990-2001/May 25
         (c) 2001 The Gale Group
  File 160:Gale Group PROMT(R)  1972-1989
         (c) 1999 The Gale Group
  File 148:Gale Group Trade & Industry DB  1976-2001/May 25
         (c)2001 The Gale Group
  File 621:Gale Group New Prod.Annou.(R)  1985-2001/May 25
         (c) 2001 The Gale Group
  File 275:Gale Group Computer DB(TM)  1983-2001/May 25
         (c) 2001 The Gale Group
  File  15:ABI/Inform(R)  1971-2001/May 26
         (c) 2001 Bell & Howell
  File   9:Business & Industry(R)  Jul/1994-2001/May 24
         (c) 2001 Resp. DB Svcs.
  File 623:Business Week  1985-2001/May W4
         (c) 2001 The McGraw-Hill Companies Inc
  File 810:Business Wire  1986-1999/Feb 28
         (c) 1999 Business Wire
  File 624:McGraw-Hill Publications  1985-2001/May 24
         (c) 2001 McGraw-Hill Co. Inc
  File 636:Gale Group Newsletter DB(TM)  1987-2001/May 25
         (c) 2001 The Gale Group
  File 813:PR Newswire  1987-1999/Apr 30
         (c) 1999 PR Newswire Association Inc
  File  20:World Reporter  1997-2001/May 26
         (c) 2001 The Dialog Corporation
  File  77:Conference Papers Index  1973-2001/May
         (c) 2001 Cambridge Sci Abs
  File  35:Dissertation Abstracts Online  1861-2001/Jun
         (c) 2001 UMI
  File 583:Gale Group Globalbase(TM)  1986-2001/May 21
         (c) 2001 The Gale Group
  File  65:Inside Conferences  1993-2001/May W3
         (c) 2001 BLDSC all rts. reserv.
 *File  65: CD=2000 and CY=2000 are not fully functioning.
Please see Help News65 for details.
  File   2:INSPEC  1969-2001/May W3
         (c) 2001 Institution of Electrical Engineers
  File 233:Internet & Personal Comp. Abs.  1981-2001/May
         (c) 2001 Info. Today Inc.
  File  99:Wilson Appl. Sci & Tech Abs  1983-2001/Apr
         (c) 2001 The HW Wilson Co.
  File 473:FINANCIAL TIMES ABSTRACTS  1998-2001/APR 02
         (c) 2001 THE NEW YORK TIMES
 *File 473: This file will not update after March 31, 2001.
It will remain on Dialog as a closed file.
  File 474:New York Times Abs  1969-2001/May 25
         (c) 2001 The New York Times
```

File 475:Wall Street Journal Abs 1973-2001/May 25
(c) 2001 The New York Times
File 634:San Jose Mercury Jun 1985-2001/May 18
(c) 2001 San Jose Mercury News
***File 634: Certain records in this file are being removed at the** publisher's request. See HELP NEWS634.
File 256:SoftBase:Reviews,Companies&Prods. 85-2001/Nov
(c)2001 Info.Sources Inc
***File 256: Please note new price changes effective May 1, 2001.** See Help Rates256 for details.
File 348:EUROPEAN PATENTS 1978-2001/May W02
(c) 2001 European Patent Office
File 349:PCT Fulltext 1983-2001/UB=20010517, UT=20010503
(c) 2001 WIPO/MicroPat
File 635:Business Dateline(R) 1985-2001/May 25
(c) 2001 Bell & Howell
File 570:Gale Group MARS(R) 1984-2001/May 25
(c) 2001 The Gale Group
File 146:Washington Post Online 1983-2001/May 21
(c) 2001 Washington Post
File 387:The Denver Post 1994-2001/May 25
(c) 2001 Denver Post
File 471:New York Times Fulltext-90 Day 2001/May 26
(c) 2001 The New York Times
File 492:Arizona Repub/Phoenix Gaz 19862001/May 25
(c) 2001 Phoenix Newspapers
File 494:St LouisPost-Dispatch 1988-2001/May 24
(c) 2001 St Louis Post-Dispatch
File 498:Detroit Free Press 1987-2001/May 24
(c) 2001 Detroit Free Press Inc.
***File 498: Certain records in this file are being removed at the** publisher's request. See HELP NEWS498.
File 630:Los Angeles Times 1993-2001/May 24
(c) 2001 Los Angeles Times
File 631:Boston Globe 1980-2001/May 25
(c) 2001 Boston Globe
File 632:Chicago Tribune 1985-2001/May 26
(c) 2001 Chicago Tribune
File 633:Phil.Inquirer 1983-2001/May 10
(c) 2001 Philadelphia Newspapers Inc
***File 633: Certain records in this file are being removed at the** publisher's request. See HELP NEWS633.
File 638:Newsday/New York Newsday 1987-2001/May 25
(c) 2001 Newsday Inc.
File 640:San Francisco Chronicle 1988-2001/May 25
(c) 2001 Chronicle Publ. Co.
File 641:Rocky Mountain News Jun 1989-2001/May 19
(c) 2001 Scripps Howard News
File 702:Miami Herald 1983-2001/May 24
(c) 2001 The Miami Herald Publishing Co.
***File 702: Certain records in this file are being removed at the** publisher's request. See HELP NEWS702.
File 703:USA Today 1989-2001/May 25
(c) 2001 USA Today
File 704:(Portland)The Oregonian 1989-2001/May 18
(c) 2001 The Oregonian
File 713:Atlanta J/Const. 1989-2001/May 24
(c) 2001 Atlanta Newspapers
File 714:(Baltimore) The Sun 1990-2001/May 24
(c) 2001 Baltimore Sun
File 715:Christian Sci.Mon. 1989-2001/May 25
(c) 2001 Christian Science Monitor
File 725:(Cleveland)Plain Dealer Aug 1991-2000/Dec 13
(c) 2000 The Plain Dealer
File 735:St. Petersburg Times 1989- 2000/Nov 01

(c) 2000 St. Petersburg Times
***File 735: This file is temporarily not updating.**
File 477:Irish Times 1999-2001/May 25
(c) 2001 Irish Times
File 710:Times/Sun.Times(London) Jun 1988-2001/May 26
(c) 2001 Times Newspapers
File 711:Independent(London) Sep 1988-2001/May 25
(c) 2001 Newspaper Publ. PLC
File 756:Daily/Sunday Telegraph 2000-2001/May 24
(c) 2001 Telegraph Group
File 757:Mirror Publications/Independent Newspapers 2000-2001/May 26
(c) 2001

| Set | Items | Description |
|---|---|---|

?

DS

| Set | Items | Description |
|---|---|---|
| S1 | 21336 | (STEGANOGRA? OR WATERMARK?) |
| S2 | 9058 | S1 AND (INTERNET OR WEB OR BROWSER) |
| S3 | 1784 | S2 AND (ADVERT? OR PHOTOP? OR (PRINT? (3W) COPY)) |
| S4 | 505 | S3 AND (PRINT? (4W) (MEDIA OR COPY OR ADVERTI?)) |
| S5 | 278 | RD S4 (unique items) |

?

Scanned all

*t s5/full/134*

**5/9/134 (Item 6 from file: 15)**
DIALOG(R)File 15:ABI/Inform(R)


01115809 97-65203
**Electronic imaging & storage guide**
Vangelova, Luba
Government Executive v27n11 PP: 1A-12A Nov 1995 CODEN: GVEXAW ISSN: 0017-2626 JRNL CODE: GOV
DOC TYPE: Journal article LANGUAGE: English LENGTH: 9 Pages
SPECIAL FEATURE: Charts
WORD COUNT: 6093

ABSTRACT: Technical barriers to widespread use of imaging have fallen, thanks to client/server arrangements, LANs, and more powerful PCs. The proliferation of imaging, with its associated benefits of improved productivity and cost reduction, is especially good news for the federal government, which relies heavily on documents in its daily operations. Several elements of imaging are discussed, including: 1. document capture, 2. storage devices, 3. image processors, 4. compression technology, 5. display systems, 6. document management and workflow, 7. search and retrieval software, and 8. systems integrators.

TEXT: Imaging Enters the Mainstream
Imaging Technology does more than save space--it also increases productivity.

Not long ago, imaging--the digitization of all types of documents--was considered an esoteric, high-end application. Soon it will be on almost everybody's desktop. Technical barriers to widespread use have fallen, thanks to client/server arrangements, local-area networks and more powerful personal computers.

Imaging technology will soon "be hard to avoid," says David McCoy, a research director at the Stamford, Conn.-based information technology advisory firm Gartner Group. The proliferation of imaging, with its associated benefits of improved productivity and cost reduction, is especially good news for the federal government, which relies heavily on documents in its daily operations.

Although more and more documents are being generated electronically, about four-fifths of the information processed in today's workplace remains in paper form. Every agency invests both personnel and equipment in the filing, retrieval, routing and storage of this paper.

Imaging technology converts paper-based information into electronic form. It then can either be stored as images or converted into machine-understood text that can be altered or searched by the computer. Any type of document is a candidate for electronic imaging: personnel records, technical drawings and medical files, for example. The documents can be in the form of printed pages, handwritten notes, photographs, drawings or maps.

Agencies using electronic document systems can expect many benefits. In the beginning, imaging was touted as a means to save storage space, but now the most sought-after benefit is increased productivity. Industry analysts say that depending on whether imaging is used alone or as part of a comprehensive process-reengineering effort, it can increase productivity by as much as several hundred percent. Cutting costs and improving service are additional incentives for using imaging technology.

Electronic documents generally take much less time to find, handle, re-file or route. And some tasks involving the same electronic document can be done

simultaneously by many employees, each handling their own "virtual" copy of the document. Labor costs also go down as fewer employees are needed for data-entry tasks; documents can be scanned and converted into electronic form by imaging hardware and software.

Electronic documents are also more easily and cheaply shared, both by co-workers and among different federal agencies. And without the need for employees to manually find and distribute a document, it can be accessed 24 hours a day. This rapid and easy access can also shorten decision-making time. Also important is that the original document remains undamaged while the virtual copies of it make the rounds.

Imaging technology is still maturing, which means that products on the market are constantly improving and costs falling. An average 20-user imaging system, including hardware and software, now costs about $8,500 per user, McCoy says. But that price has been dropping dramatically. McCoy expects costs to continue dropping about 15 to 20 percent per year through the year 2000. Low-end imaging products, such as machines to scan business cards, also are available. A single-user imaging system for the PC called Discovery Edition, made by **Watermark** Software Inc. of Burlington, Mass., costs just $149.

A basic imaging system is composed of a standard workstation as well as a scanner, a storage device, an image-display monitor and image-processing, compression and search-and-retrieval software. But while this type of setup can benefit an office just by eliminating paper, using imaging technology as an electronic filing cabinet is not using it to its best advantage. That's why it is often combined with workflow and/or document management software. Creating an effective imaging system usually requires using many vendors and products, so systems integrators are frequently enlisted.

As is true for most areas of information technology, but especially for still-maturing markets, the pace of technological change in the imaging field is very fast. Faster, more accurate and more sophisticated products are constantly making their debuts. Although users once had to make a large investment in new hardware before installing an imaging system, vendors today are offering many more solutions that fit into existing computing environments. Therefore, the software and services portions of the imaging market are growing, while expenditures on hardware are falling.

Another reason for this trend is that the functions of several hardware components now sometimes can be found in one integrated product: For example, Canon's Canofile 510 Desktop Electronic Filing System uses a scanning/storage/display unit. Some hybrid products can even handle both digital images and micrographics, which is useful for agencies with large investments in that technology.

Many areas of imaging are plagued by a lack of standards. The incompatibility among some vendors' products and the uncertainty of finding support for a particular product in the future gives potential users reason for pause. But there is now a movement toward open architectures and standards. For instance, the Document Management Alliance, a new coalition formed this year, plans to introduce standards for document-management software.

Alliances between vendors also are spurring the development of common formats. For example, Kodak will use Wang's software in its document-management software solutions, and the two companies intend to develop common document-imaging architectures. Meanwhile, under a separate agreement with Microsoft, Wang will provide desktop imaging software to be bundled with future releases of Microsoft's Windows 95 and Windows NT. James E. Breuer, senior vice president of marketing for the Association for Information and Image Management International (AIIM), an industry trade group, calls this "probably the biggest announcement in imaging this year."

The Microsoft-Wang agreement, as well as Lotus's decision to bundle image viewing, processing and transmission software with its Lotus Notes and cc:Mail, will move imaging into the mainstream by allowing it to migrate from the mainframe to the desktop.

This move means imaging will also be used more in applications other than transaction processing. For instance, more individual users will begin to scan documents for purposes such as word processing.

The imaging market, which only came into its own in the mid-1980s, almost doubled its revenues between 1995 and 1999, according to BIS Strategic Decisions Inc., a Norwell, Mass.-based market research firm. The federal government, which has undertaken both larger imaging projects and many small-scale pilot projects, spent $700 million on imaging technology in fiscal 1994, according to INPUT, a Vienna, Va.-based market research firm. The company projects that this total will grow to almost $2 billion by fiscal 1999. According to a recent survey by AIIM, the federal government ranks fourth on the list of industries in which imaging technology has penetrated the most deeply.

Some of the biggest, best-known federal imaging projects under way are the IRS's Document Processing System, which will image and store tax returns and related correspondence; the FBI's Integrated Automated Fingerprint Identification System, which will use imaged fingerprints to speed the identification of criminal suspects; and the Patent and Trademark Office's Automated Patent System, which will result in an image-based storage and retrieval system for patent documents.

But the high start-up cost of a comprehensive imaging system still makes some managers cringe. Partly for this reason, management approval remains a common barrier to implementing imaging technology. And legacy systems in which departments have heavily invested are not always compatible with the new technology.

Because of paper's convenience and portability advantages, there will never be a totally paperless office. Nevertheless, imaging use will continue to rise. In the future, some documents will be almost unrecognizable in comparison with today's. Users will be able to attach multi-media notes, such as voice messages or full-motion video, to documents. And imaging will be a part of so many different applications that it will eventually be impossible to separate the imaging components from the rest of the information technology.

Document Capture

Printed material can be converted to digital form in-house or off-site.

The first task in imaging is to convert printed materials into digital form. This can be done either in-house using a scanner (or digital camera in some cases), or off-site by a company specializing in document conversion. Document capture is generally the most expensive imaging system component if image processing is figured into the cost; in fact, image-processing is now often integrated into document-capture products and services. (Image processing is covered on Page 8A of this guide).

Scanner prices have dropped dramatically, and some now sell for as little as several hundred dollars. In addition to scanning documents, some multi-function products now **print** , **copy** and fax as well. One scanner is distinguished from another primarily by the resolution at which it reads documents and the number of pages in can process per minute. Top speeds are about 120 pages per minute, although there are a few even faster scanners. Different scanners also work better with different document sizes and forms (for example, maps versus text pages).

Scanners available today include: basic hand-held models that: a person

drags across the document to be scanned; sheet-fed models that resemble fax machines; flatbed scanners that resemble copiers, in which the original is placed on a flat, glass plate; drum scanners, in which the original is placed on a rapidly rotating cylindrical drum; slide and transparency scanners; and high-speed scanners that can handle microfilm or large-format documents such as engineering drawings.

Many scanners come with automatic document feeders and offer color scanning and the ability to scan two-sided pages. Some, such as Bell & Howell Document Management Products Company's FD4400 Film Digitizing System, also work with microfilm, converting film images into digital form. Kodak's Imagelink scanner/microimager 990D converts paper-based documents to both digital and micrographic media. Bell & Howell's Copiscan II line of scanners now offers an imprinter option that can stamp such things as date and time onto scanned documents.

Digital cameras are ideal for remote-image acquisition, such as might be necessary during military operations, when images captured on location need to be transferred back to headquarters for analysis. One product that can do this is Eastman Kodak Company's Megapixel Imaging Technology camera system, which can capture up to 30 frames per second. Digital camera prices have dropped significantly since the technology was first made available. Kodak's Digital Camera 40, the company's first point-and-shoot digital camera, now sells for less than $1,000. Images stored in digital cameras can be downloaded to a personal computer.

Sometimes agencies have too many documents requiring transfer into electronic format to effectively handle the document-capture step in-house. Such projects can be contracted out to companies that specialize in document conversion. The conversion can be done either at one of the contractor's centers or at the client's offices, which could be useful when classified data is involved. The outsourcing of data conversion is most often needed for backfile conversions, in which historical records are converted to digital form. Conversion firms usually index the data once it is scanned in, and they may perform other image processing functions.

Backfile conversion is not always necessary. It is very expensive, and may not be worthwhile unless users anticipate a need for random access to historical files or hope to perform efficient searches on historical data. A common conversion strategy known as "day-one-forward" designates all future files for scanning and leaves historical records as they are. Only once a historical record is required for a task is it scanned in.

One company that provides document conversion services is Unicor, otherwise known as Federal Prison Industries. Part of the U.S. Department of Justice, the Washington-based company employs inmates. It has digitized archival books, films and photos for the Library of Congress.

I-NET of Bethesa, Md., is building a system to be installed at the U.S. Navy Annex in Arlington, Va., to digitize and index 65 million Navy personnel records that currently reside on microfiche. I-NET is also working on a $12 million contract to convert all of the Federal Aviation Administration's Civil Aviation Registry's historical files--aircraft. records and airman certification records--from paper, microfilm and microfiche to digital images. During the conversion process, each document will be assigned a bar code ID to allow registry employees to continue to access the records.

West Coast Information Systems Inc. (also known as WESCO) of Walnut Creek, Calif., has already converted aperture cards containing architectural and construction engineering information into electronic images for the National Institutes of Health's Design and Construction Branch. Now it is converting the branch's paper-based drawings into digital images, using aperture cards as an intermediary step in the process.

Storage Devices

Imagining storage systems house a lot of data in very little space.

One of the primary advantages of electronic imaging is that storage options consume much less physical space than do paper documents or even microfiche and microfilm. The two most popular choices for storing electronic images are magnetic media and optical disks. Both of these also offer generally faster retrieval times than paper or micrography. The prime considerations in choosing an imaging storage system are the type of data, the required speed of access to the data and the cost of the storage.

Magnetic storage devices--such as a floppy disk or the hard drive on a network server or desktop computer-offer the fastest access to data but do not hold much. However, recent technology advances allow magnetic storage to adequately store all of the data required for smaller imaging systems. It can also serve to cache the most frequently used data from a larger system.

Optical systems, which offer more space and greater durability than magnetic media, depend on lasers to store and read data. These systems can store from about 128 megabytes up to more than 1.3 gigabytes of information. Their cost per megabyte is lower than for magnetic media but higher than for slower-access storage options such as tape. Fortunately, optical storage prices continue to drop.

One of optical storage media's most attractive features is that it offers random access to data (as opposed to tapes, which store data sequentially and require the user to travel past all previously recorded data in order to access a particular snippet of information). Also, because they can store a lot of data in very little space, optical media can be used to archive data on-site, which makes historical data easier to access.

Optical disks are usually either 3 1/2 inches, 5 1/4 inches, 12 inches or 14 inches in diameter. The optical storage category includes WORM (Write Once, Read Many), CD-ROM (Compact Disk-Read Only Memory), magneto-optical and now CD-Recordable (Compact Disk-Recordable) systems. Each has distinct advantages and drawbacks.

CD-ROMs, although they don't provide as much room as some other options, cost very little to produce in quantity and therefore may be useful for data that needs to be distributed widely. They also guarantee that the images stored on them are unalterable, which is important for some applications, especially in the financial and legal arenas. Industry analysts expect to see much higher-capacity CD-ROMs entering the market, and they also expect to see CD-ROMs that offer rewritable capabilities.

WORM disks are also unalterable, and they store more information than CD-ROMs, although they cost more. One recent development has married the economy of CD-ROMs with convenience. CD-Recordable devices, which entered the mainstream imaging market in 1995, allow users to produce CD-ROMs at their desktops. Micro Design International Inc. of Winter Park, Fla., makes the ExpressWriter, one of the products that offers CD-Recordable technology.

Plasmon Data of Milpitas, Calif., has a unit it calls a PD/CD-ROM drive, which combines high-capacity rewritable optical storage with the advantages of a high-speed CD-ROM player. Other optical storage vendors include Nikon, which offers high-capacity magneto-optical disk drives, and Philips LMS.

Last year, the Optical Storage Technology Association endorsed a new optical file system standard to allow optical disks to be used across platforms and to be compatible with other vendors' products.

For storing information that may need occasional updating, agencies can use

rewritable optical systems, such as magneto-optical disks. For offices that require storage flexibility, vendors such as Hewlett-Packard Co. and Sony Electronics Inc. now offer multifunction optical drives that can handle either WORM or rewritable disks.

There is an inverse relationship between the cost per megabyte and the access speed of the various storage media. Some offices have turned to hierarchical storage management, a system whose goal is to store data on the most efficient combination of different media. The most-often accessed data is stored online on magnetic media; once it is needed less frequently, the data is stored near-line on optical storage; and then it is archived on tape, such as digital audio tape or conventional 1/4-inch or eight-millimeter analog tape cartridges. Copies of important data can also be stored separately off-site. Hierarchical systems, such as Micro Design International Inc.'s EZ-Express, move data between media as needed, but in a way that is transparent to users.

For imaging systems that require more capacity than one optical disk or magnetic tape can accommodate, agencies can use jukeboxes, which provide automated access to all stored data by using robotic arms to select and activate the needed disk or tape. Jukeboxes are becoming faster and more reliable, but occasionally logjams can occur when too many people try to use them at once.

Some jukeboxes are now equipped to handle both optical disks and magnetic tape. For example, ATG Cygnet Inc. of San Jose, Calif., makes the ASM/Cygnature jukebox, which stores tape, magneto-optical disks and WORM disks.

Image Processors

Software can transform handwritten documents into computer-readable text. Sometimes it is necessary o manipulate scanned images before they are stored. Image processors run the gamut from soft ware that converts text or even handwriting imbedded in images into computer-readable information, to software that enhances document quality and readability.

A computer cannot recognize words within an image as words unless they are converted into the same format in which computers store text. The software that accomplishes this is called Optical Character Recognition (OCR). If the text in question is handwritten, then more advanced techniques are called for; this type of software is known as Intelligent Character Recognition (ICR). OCR/ICR-processed documents can be searched, edited and manipulated just like any other electronic documents.

Optical Mark Recognition software is another type of image processor that transforms marks--such as the check marks or Xs often found on forms--into computer-recognizable information. The Census Bureau will issue a request-for-proposals next June for the Census Data Capture System 2000, which is expected to make use of such software to automatically process census returns.

The accuracy of some OCR software is very high, but there is always room for improvement. Advanced techniques now used to improve the accuracy of OCR/ICR include analyzing the context in which words appear and consulting application-specific dictionaries. Some recent OCR software releases, such as Teleform 4.0 from Cardiff Software Inc. of Solana Beach, Calif., use a voting scheme in which several OCR techniques are applied and their results compared. Teleform 4.0 also applies fuzzy logic, an artificial intelligence technique that allows the software to formulate an educated guess when the character it's reading is unclear.

OCR/ICR software is designed to process either pages, which contain primarily free-form text, or forms, in which textual information is confined to predetermined zones. In forms processing, the picture of the

form is usually removed and the text stored separately for more efficient use of space. It can then be recombined with the form whenever necessary.

If more than one type of form is to be processed, the forms must usually be sorted manually first. But some imaging products, such as Kodak's ICMS family of scanners, have built-in image processing software with forms recognition capability that makes preliminary sorting unnecessary.

The Internal Revenue Service's Document Processing System will use software that recognizes each form and generates a new image by subtracting the image of the form itself. It will also use ICR to convert handwritten information. Because accuracy is so important to the IRS, the image-processing software will apply several processing algorithms, but it will not offer a guess when it can't read a character; instead, these forms will be passed to humans to process.

Vendors selling OCR software include Caere Corp. of Los Gatos, Calif., Calera Recognition Systems of Sunnyvale, Calif., and Xerox Corp. Vendors of ICR software include AEG Corp. of Bethesda, Md.; ComCom Systems of Clearwater, Fla., Cognitronics Corp. of Danbury, Conn.; and Symbus Technology Inc. of Waltham, Mass. AEG's HL-ICR 6160, for instance, can read images containing hand-printed uppercase letters.

Adobe has introduced a new method of image processing with its Acrobat Capture product. It's a page recognition program that works with many scanners and uses OCR and other processing techniques to convert paper documents into a computer-searchable and -indexable format. Its unique feature is its ability to recognize text, page layout, fonts and graphics and to turn the document into an electronic file that looks exactly like the printed original. It both maintains the formatting that most OCR products lose and allows access to the content of the document, which is not possible with an un-processed document. The Department of Defense evaluated Acrobat Capture during its development.
Other types of software exist to make scanned documents sharper and easier to read; to "clean" documents by removing stray marks or stains; to re-align an image; and to automatically crop an image. When such software is used before OCR processing, it can increase the OCR accuracy. Sequoia Data Corp. of Burlingame, Calif., claims its ScanFix software--which straightens crooked lines, removes rules and boxes and cleans up marks like coffee rings--an increase OCR accuracy by 50 percent.

Post-conversion processing software can tag data and load it into various databases or perform functions such as checking for duplicates. Indexing documents is another important function that ensures the documents will be easy to find once they are in the computer. Documents can be indexed either by a few key fields or by the full text. Pictures can even be indexed: Excalibur Technologies Corp. of McLean, Va., makes the XRS Print Recognition Module, which automatically indexes fingerprints.

Image processing software often requires considerable computing power, which has only just become common in personal computers. These functions often are bundled with comprehensive document imaging systems or with stand-alone scanners, but they also can be bought as separate packages.

Compression Technology
Compressed images save space--and therefore, money.

Access and retrieval times decrease if a captured image is compressed to take up less storage space. The data can also thus be stored more inexpensively. In some cases, less-necessary parts of the data are lost in the compression step, but in others, all the data remains in place, just in different form. Before an image is viewed or printed, it is decompressed, but this step is much faster and cheaper than compression.

Using one standard compression technique, still images can be squeezed into

one-fourth the space. But compression ratios can be increased to about 50-to-1. or even more when using sophisticated mathematical formulas to describe image components.

Compression technology takes the form of either hardware or software. TMS Inc. of Stillwater, Okla., developed one of the first all-software image compression/decompression toolkits. The level to which data can be compressed depends on both the condition of the document and the way in which it is laid out.

Compression technology decreases the size of an imaging file by eliminating unnecessary information, such as gaps, empty data fields or redundancies. For example, in a video file, if an object such as a house appears in every frame, compression technology can encode it once and concentrate on the information that changes from frame to frame. Also, white areas on a printed page can be saved as chunks, instead of devoting a piece of storage space to each tiny dot.

Some companies that offer compression technology are: TMSInc., which includes it in a developer toolkit called ViewDirector, and TASC (The Analytic Sciences Corp.) of Reading, Mass.

Display Systems

High-tech monitors enhance electronic-document viewing.

Although images can be viewed on standard computer monitors, it can be useful to invest in high-resolution, large-screen monitors, especially if increasing productivity is one of the primary aims of installing the system. Large screens, from 17 inches to 21 inches, are easier to read, and they are more useful for showing two documents side-by-side. They also enable employees to view an entire 8 1/2 x 11-inch document without having to scroll.
A high-resolution monitor is necessary if a user wants to work with an on-screen document whose resolution approaches that of a paper document. Resolution is determined by the number of pixels--the illuminable dots that comprise an electronic image--per inch. The higher the number of pixels used to create a character or graphic on the screen, the clearer and easier to read it is. The highest-resolution display systems available on the market feature several hundred dots per inch.

Anti-glare screens also make the documents on display easier to read. Grayscale monitors offer a higher image quality, but whether the display monochrome, grayscale or color is largely a matter of preference and budget, unless there is a functional reason to require color.

One display system designed to enhance electronic document viewing is the "Color 20/70" from Cornerstone Imaging of San : Jose, Calif. It features a 20-inch, flicker-free, high-resolution display with anti-glare coating. Document Technologies Inc. of Sunnyvale, Calif., incorporates some image-processing software into its EasyRead 240 grayscale display. It smoothes and tunes each character to make the text on the screen even easier to read.

Document Management and Workflow
Offices must rethink their business processes before installing workflow software.

Serving as electronic filing cabinets, "the majority of imaging applications are file-and-retrieve applications, which is not the highest and best use of imaging," says James E. Breuer, senior vice president of marketing for the Association for Information and Image Management International (AIIM). Instead, imaging should be considered an enabling technology for improving business processes and making them more efficient. One of imaging's most important features is that, combined with workflow

tools, it can "take practices that were sequential and make them simultaneous," Breuer says. This requires rethinking the entire business processes.

So it is not surprising that imaging is often used in conjunction with software that automates business processes. Today, workflow software and its less-comprehensive sister technology, document management software, are often sold as a package dear with imaging software. Demand for workflow tools is growing: a recent user survey conducted by AIIM showed that 62 percent of users intend to use them in the future.

A workflow system, which is often closely linked to business process reengineering efforts, consists of hardware, software or both. The system is programmed with a particular businesses' routing procedures and processing rules. This allows it to automatically route documents to the employee or department to which they must go to complete a transaction. An employee's daily workload can therefore await her when she signs on in the morning. A workflow system can also send documents to databases for storage. Workflow provides comprehensive reporting functions, and it integrates applications such as word processing, e-mail and fax.

The potential benefits are high, but installing a workflow system requires a great deal of up-front work to re-think the business process. It is also costly. The average per-user cost of a typical entry-level workflow system is about $3,500.

The workflow market is still relatively new; almost half of the products available today have been introduced within the last two years. The Workflow Management Coalition is working to set standards. And Wang and Microsoft are cooperating to define open work management interfaces. But in e next several years, there probably will continue to be volatility in the market as some vendors fall by the wayside. Analysts expect to see workflow capability built into a wide range of application packages in the future.
FileNet Corporation of Costa Mesa, Calif., which makes WorkFlo, is one of the most recognized providers of workflow software. Other vendors are Keyfile Corp. of Nashua, N.H.; IBM; and Staffware Corp. of Wellesley, Mass. Wang and Kodak plan to offer an integrated imaging/workflow portfolio. Wang's OPEN/workflow product gives users point-and-click capability to construct a graphical model of the work process they follow and to define its rules. OPEN/workflow also takes into account manual tasks as it automatically controls the business procedures. Computron Software Inc. of Rutherford, N.J., sells an open-architecture workflow product it calls Computron Workflow, which can be integrated with other applications used to run a business.

Document-management software is often used in conjunction with imaging and can also be integrated into an imaging/workflow system to realize even greater productivity gains. Whereas workflow software focuses on automating business procedures, document-management software is a management tool for electronic documents, automatically creating, managing and organizing them. It can be useful in improving accountability and helping agencies comply with federal policies regarding document storage. For example, it can track document revisions. It is most useful when it is applied to all of the documents used in a business process.
Examples of document-management software are: MI sup 3 MS from Minolta; PageKeeper from Caere; PageNet from PageNet Inc. of Mississauga, Ontario, Canada; CabiNet for Windows from CabiNet Systems Inc. of Mahwah, N.J.; and ImageFast from Compusearch Software Systems Inc. of McLean, Va.

Workflow software also can be used to measure the effectiveness of an electronic document management system. For instance, it can keep track of the number of documents being processed in a given time frame.

One federal department taking advantage of both imaging and document-management software is the Department of Veterans Affairs. In

September, the department awarded Oracle Complex Systems a $34 million contract, part of which will go toward modernizing the Veterans Benefits Administration's claims process using an image-based document-management system.

The Internal Revenue Service is using workflow as part of its Service Center Recognition Image Processing System. The workflow system automatically routes forms and monitors system performance and resource use.

Search and Retrieval Software
Systems index information by date, employee name, key words and more.

Stored documents can be retrieved easily if they are properly indexed. Search-and-retrieval software can handle both structured data--documents such as forms, which contain information in predictable fields-and unstructured data-collections of different types of documents, such as articles, forms and reports.

Structured data, common in transaction-based environments, can usually be searched by field. For instance, the user can request a document by date, employee name or other predetermined category. Unstructured data, which is commonly found in intelligence, regulatory and library applications, is often best located using full-text search. In this case, the user enters a few key words, and the search-and-retrieval software responds with a list of all documents that contain that combination of words.

Some of the most advanced systems can retrieve documents with relevant information, even if the specific key words did not appear within them. Search techniques range from basic boolean (which can find direct matches to words, or groups of words linked with "and" or "or") to the fuzzy logic type of artificial intelligence.
Relational database management systems often play a role in search-and-retrieval software. In such a database, imaging data is stored in arrays of rows and columns, which facilitates finding all data which falls within a certain range--for example, employees who earn between $30,000 and $35,000 per year.

Off-the-shelf retrieval software for a personal computer now costs between $100 and $200; for larger systems the price can run into the tens of thousands of dollars.

Search-and-retrieval software is often integrated into comprehensive imaging and document-management systems. But it can also be bought separately. Excalibur Technologies Corp. of McLean, Va., makes the popular EFS document management and retrieval software. Products such as Computron's COOL (Computer Output OnLine) incorporate retrieval software that automatically recombines retrieved data with an image of the form on which it appears. (The two are stored separately for efficiency reasons.)

A burgeoning segment of the search-and-retrieval market is in the area of full-text **Internet** search packages for the World Wide **Web** , which is essentially a large repository of publicly available information. One such product is Latitude from Open Text of Waterloo, Ontario, Canada.
Systems Integrators

Depending on the size and scope of an imaging system, products from up to several dozen vendors may be needed. And within each imaging product category, there is a wide range of functionality between different products. Finding the right imaging products and combining them in a manner in which they can work together is a challenge for all but the most basic applications. Also, a lot of time can be expended gathering competitive bids, managing the procurement process and evaluating the products. Agencies often find the solution is to outsource the project to a systems integrator.

Integrators "are experts in the field," says Michael J. Cocchiola, director of the Defense Printing Service in Washington. His office is managing the $38 million Automated Document Management And Publishing System (ADMAPS). It awarded a systems-integration contract to Eastman Kodak Company in 1991 for three years with an option to extend for two more; the contract is currently in the follow-on option stage. "We can't keep up with the technology," Cocchiola says. "Using an integrator saves us time and money."

The DoD ADMAPS project is a printing, imaging, document management and storage "erector set" intended to reduce the cycle time and cost of storing and retrieving DoD technical documents. Kodak, as the prime contractor and systems integrator, has assembled systems that accept information in almost every format and type. Users are able to then edit and combine data from different document sources and output data in a wide range of formats. Each specific configuration may be different depending on the site's unique needs.

Systems integrators first determine the customer's requirements in such areas as the number of users, required access speed, the number of documents to be processed and storage needs. Their job is to then find the best-fit product in each needed category and to combine the products in a way that meets the user's needs while maintaining an ease of use. Often, systems integrators will also arrange training to acquaint the staff with the new technology.

Some projects begin with a small-scale pilot system, as did the Internal Revenue Service's Document Processing System, one of the federal government's largest imaging projects. For smaller or shorter-term projects, a pilot is not always necessary or able to fit into the schedule.

A factor critical to a project's success is communication between the integrator and the department's management. This should continue throughout the project life-cycle to ensure that the needs remain the same and that the system under development will meet them. It is also useful to involve end users in design and development.

Imaging is frequently used as a technology enabler in business process reengineering, and a systems integrator who specializes in BPR can be invaluable. It is especially important, then, to find an integrator who already has some experience with the type of process to be converted.

Some integrators with federal government imaging experience include PRC Inc. and BTG Inc., both of McLean, Va., and Kodak. Loral Federal Systems-Gaithersburg of Gaithersburg, Md., is the prime contractor on the IRS Document Processing System. The system is part of the IRS's 10-year, $8 billion Tax System Modernization program, which is aiming to make the processing of the hundreds of millions of tax forms and letters the IRS receives each year more efficient and economical.

The IRS explained its needs to Loral and left the equipment selection up to the integrators. The new system will use electronic scanning, image processing and storage to handle this information load. Loral has assembled a system using a Kodak scanner, IBM servers, Cornerstone high-resolution monitors and Kodak optical disks, among other products.
Upcoming large imaging projects include a remote sensing and image processing system for the U.S. Forest Service. Federal Sources Inc., a McLean, Va., market research firm, expects the department to issue a request for information in early 1995 and estimates that the contract may eventually add up to about $50 million. The Army is expected to request a follow-on recompeting bid on its Medical Diagnostic Imaging System, on which Loral has been the prime contractor. Federal Sources estimates that the value of the new contract could be as much as $300 million.

And the contract for one of the federal government's biggest imaging

projects, the Patent and Trademark Office's Automated Patent System, will expire next September. The system scans, stores and retrieves millions of domestic and foreign patents that need to be accessed by government examiners, inventors and manufacturers. PTO is expected to want to continue to receive integration services for the system after the current contract, held by PRC, expires. A new draft request-for-proposals has been issued, and Federal Sources expects the contract to amount to about $500 million.


GEOGRAPHIC NAMES: US

DESCRIPTORS: Image processing systems; Information storage; Technological change; Productivity; Cost control
CLASSIFICATION CODES: 5230 (CN=Computer hardware); 5240 (CN=Software & systems); 9190 (CN=United States)
?